



STIC Search Report

EIC 2100

STIC Database Tracking Number: 201167

TO: Cam-Linh T Nguyen

Location:

Art Unit: 2161

Monday, September 11, 2006

Case Serial Number: 9801340

From: Lance Sealey

Location: EIC 2100

RND-4B11

Phone: 571-272-8666

Lance.Sealey@uspto.gov

Search Notes

Dear Cam-Linh,

These were the closest results I could find.
Please let me know if you have any questions.

Lance

STIC Search Results Feedback Form

EIC 2100

Questions about the scope or the results of the search? Contact *the EIC searcher or contact:*

Alyson Dill, EIC 2100 Team Leader
272-3527, RND 4B28

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 2133

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(Journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/EIC2100 RND, 4B28

(FAST AND FOCUSED SEARCH)

Access DB# 201167

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: Nguyen, Cam Linh Examiner #: 78921 Date: 9/11/06
Art Unit: 2161 Phone Number: 2-4024 Serial Number: 091801, 340
Mail Box and Bldg/Room Location: 3A35 Results Format Preferred (circle) PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: System & method for Identifying word patterns in text

Inventors (please provide full names): Simpson, Don
Uney, Robert

Earliest Priority Filing Date: 3/7/01

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

- word pattern in (query or "stream of text")
- semantic network

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>LANCE SEALEY</u>	NA Sequence (#) _____	STN _____
Searcher Phone #: <u>2-8666</u>	AA Sequence (#) _____	Dialog _____
Searcher Location: <u>RND 4B11</u>	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>9/11/06</u>	Bibliographic _____	Dr.Link _____
Date Completed: <u>9/11/06</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: <u>99</u>	Fulltext _____	Sequence Systems _____
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: <u>122</u>	Other _____	Other (specify) _____

EAST Search History

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	970	semantic near3 network\$3	US-PGPUB; USPAT; EPO; JPO; DERWENT ; IBM_TDB	OR	OFF	2006/09/11 10:29
L2	10	(semantic near3 network\$3) same (word\$3 near2 pattern\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT ; IBM_TDB	OR	OFF	2006/09/11 10:30
L3	21	(semantic with network\$3) same (word\$3 with pattern\$1)	US-PGPUB; USPAT; EPO; JPO; DERWENT ; IBM_TDB	OR	OFF	2006/09/11 11:01
L4	11	3 not 2	US-PGPUB; USPAT; EPO; JPO; DERWENT ; IBM_TDB	OR	OFF	2006/09/11 10:57
L5	16255	(707/1 or 707/2 or 707/3 or 707/4 or 707/6 or 707/5 or 706/55).ccls.	US-PGPUB; USPAT; EPO; JPO; DERWENT ; IBM_TDB	OR	OFF	2006/09/11 11:00
L6	33	(semantic with network\$3) and (word\$3 with pattern\$1) and 5	US-PGPUB; USPAT; EPO; JPO; DERWENT ; IBM_TDB	OR	OFF	2006/09/11 11:02
L7	29	6 not 3	US-PGPUB; USPAT; EPO; JPO; DERWENT ; IBM_TDB	OR	OFF	2006/09/11 11:02

Set	Items	Description
S1	2	AU=((SIMPSON D? OR SIMPSON, D?) AND (USEY R? OR USEY, R?))
S2	0	(AU=(SIMPSON D? OR SIMPSON, D? OR USEY R? OR USEY, R?) AND SEMANTIC) NOT (AD=(20000307:20030307) OR AD=(20030308:2006091-1))

? show files

File 347:JAPIO Dec 1976-2005/Dec(Updated 060404)

(c) 2006 JPO & JAPIO

File 350:Derwent WPIX 1963-2006/UD=200657

(c) 2006 The Thomson Corporation

1/5/1 (Item 1 from file: 350)
DIALOG(R)File 350:Derwent WPIX
(c) 2006 The Thomson Corporation. All rts. reserv.

0013650250 - Drawing available

WPI ACC NO: 2003-746275/

XRPX Acc No: N2003-597957

Object identifying method for Internet applications, involves receiving incoming stream of text comprised of words, identifying word patterns in text stream and referencing objects corresponding to word pattern

Patent Assignee: SIMPSON D M (SIMP-I); USEY R W (USEY-I)

Inventor: SIMPSON D M ; USEY R W

Patent Family (1 patents, 1 countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update
US 20030167276	A1	20030904	US 2001801340	A	20010307	200370 B

Priority Applications (no., kind, date): US 2001801340 A 20010307

Patent Details

Number	Kind	Lan	Pg	Dwg	Filing Notes
US 20030167276	A1	EN	15	7	

Alerting Abstract US A1

NOVELTY - The method involves receiving an incoming text stream comprised of words by a software text analysis module. A semantic network is consulted to identify the word patterns located at nodes in the network where the patterns correspond to objects of the network in the stream of text in a single examination of each word and a known object corresponding to the word pattern of the network is then referenced.

DESCRIPTION - An INDEPENDENT CLAIM is also included for a system of identifying objects in a data stream.

USE - Used for identifying word pattern in Internet applications.

ADVANTAGE - The word patterns are identified in a quick and efficient manner to improve research efforts and enhance the abilities of users to profitably use the Internet.

DESCRIPTION OF DRAWINGS - The drawing shows a schematic block diagram illustrating a word pattern identification module used in object identifying method.

- 200 Word pattern identification module
- 202 Semantic network generation module
- 204 Text analysis module
- 206 Base interface
- 208 Object parser
- 214 Word pattern placement module
- 216 Node linking module

Title Terms/Index Terms/Additional Words: OBJECT; IDENTIFY; METHOD; APPLY; RECEIVE; INCOMING; STREAM; TEXT; COMPRISE; WORD; PATTERN; REFERENCE; CORRESPOND

Class Codes

International Classification (Main): G06F-007/00

US Classification, Issued: 707102000

File Segment: EPI;

DWPI Class: T01

Manual Codes (EPI/S-X): T01-F07; T01-J16A; T01-J16C3; T01-N03A2

Set	Items	Description
S1	699601	SENTENCE? ? OR PARAGRAPH? ? OR DOCUMENT? ? OR (STREAM? ? (3-N)TEXT) OR HYPERTEXT? ? OR WEBPAGE? ? OR HOMEPAGE? ? OR PAGE? ?
S2	923589	WORD? ? OR TOKEN? ? OR TERM? ? OR STRING? ?
S3	78934	(TOKENI???? OR PARS??? OR ANALY???? OR EVALUAT??? OR EXAMI-N???? OR DIVID??? OR DIVIS??? OR SEPARAT??? OR SPLIT???? OR -GROUP??? OR SECTION??? OR SEGMENT???? OR SUBDIVI???? OR -PARTIT???? OR COMPARTMENT???? OR SUBSET? ? OR SUB() (SET? ? OR TYPE? ? OR
S4	2428	S3 (5N) S2
S5	661498	PATTERN? ? OR SYNTA???? OR TEMPLATE? ? OR PHRASE? ? OR R-ELATIONSHIP? ?
S6	21103	S2 (3N) S5
S7	3217	(CONSTRUCT??? OR BUILD??? OR CREAT??? OR MAKE OR MAKING OR MADE OR DEVELOP??? OR RECREAT??? OR ASSEMBL??? OR GENERAT??? -OR NEW OR FRESH OR " () SCRATCH OR DERIV???? OR PRODUC??? OR -FORM??? OR SYNTHESI???? OR MANUFACTUR???) (5N) S6 FROM"
S8	3217	S7 (5N) S2
S9	32919	(EACH OR ALL OR EVERY) (2W) S2
S10	1962	(SEARCH??? OR RETRIEV??? OR QUER???? OR MATCH??? OR FIND??? OR TARGET???) (5N) S9
S11	126	S10 (5N) S6
S12	485	SEMANTIC (3N) (NETWORK? ? OR GRAPH? ? OR TREE? ? OR NET? ?)
S13	0	S11 (5N) S12
S14	0	S11 (10N) S12
S15	7	S11 AND S12
S16	4	S4 AND S8 AND S15
S17	0	(S12 AND IC=(G06F-007/00)) NOT (S16 OR PD=(20000307:200303-07) OR PD=(20030308:20060911))
S18	0	(S12 AND S5/TI AND (SEARCH??? OR RETRIEV??? OR QUER???? OR MATCH??? OR FIND??? OR TARGET???) /TI) NOT (S16 OR PD=(2000030-7:20030307) OR PD=(20030308:20060911))
? show files		
File 348:EUROPEAN PATENTS 1978-2006/ 200636		
(c) 2006 European Patent Office		
File 349:PCT FULLTEXT 1979-2006/UB=20060831UT=20060824		
(c) 2006 WIPO/Thomson		
?		

690 TOKENI????
 28422 PARS???
 522444 ANALY????
 303189 EVALUAT???
 1927722 EXAMIN?????
 454280 DIVID???
 221566 DIVIS???
 1431280 SEPARAT???
 145868 SPLIT????
 828414 GROUP???
 986697 SECTION???
 258136 SEGMENT?????
 45001 SUBDIVI??????
 97078 PARTIT?????
 87700 COMPARTMENT?????
 73592 SUBSET? ?
 454904 SUB
 1016838 SET? ?
 1382766 TYPE? ?
 806850 GROUP? ?
 16749 SUB(W) ((SET? ? OR TYPE? ?) OR GROUP? ?)
 99788 CLEAV???
 28146 SEGREGAT???
 7577 SUBSECTION???
 699601 S1
 S3 78934 (TOKENI???? OR PARS??? OR ANALY???? OR EVALUAT??? OR
 EXAMIN????? OR DIVID??? OR DIVIS??? OR SEPARAT??? OR
 SPLIT???? OR GROUP??? OR SECTION??? OR SEGMENT????? OR
 SUBDIVI??????? OR PARTIT????? OR COMPARTMENT????? OR
 SUBSET? ? OR SUB() (SET? ? OR TYPE? ? OR GROUP? ?) OR
 CLEAV??? OR SEGREGAT??? OR SUBSECTION???) (5N) S1

685168 CONSTRUCT???

227575 BUILD???

578507 CREAT???

632422 MAKE

579755 MAKING

1323360 MADE

534114 DEVELOP???

12303 RECREAT???

564467 ASSEMBL???

912007 GENERAT???

800244 NEW

114452 FRESH

811 FROM

20230 SCRATCH

0 FROM (W) SCRATCH

591705 DERIV?????

1341231 PRODUC???

2119101 FORM???

281171 SYNTHESI????

739390 MANUFACTUR???

21103 S6

S7 3217 (CONSTRUCT???

OR BUILD???

OR CREAT???

OR MAKE OR MAKING

OR MADE OR DEVELOP???

OR RECREAT???

OR ASSEMBL???

OR

GENERAT???

OR NEW OR FRESH OR "FROM" () SCRATCH OR

DERIV?????

OR PRODUC???

OR FORM???

OR SYNTHESI????

OR

MANUFACTUR???) (5N) S6

16/3,K/2 (Item 2 from file: 348)
DIALOG(R) File 348:EUROPEAN PATENTS
(c). 2006 European Patent Office. All rts. reserv.

00602441

Method for resolution of natural-language queries against full-text
databases

Verfahren, um natursprachliche Abfragen von Textdatenbanken zu lösen

Procede pour resoudre des demandes en langage naturel dans des bases de
donnees de textes

PATENT ASSIGNEE:

CONQUEST SOFTWARE INC., (1713100), 9700 Patuxent Woods Drive, Suite 140,
Columbia, Maryland MD-21046, (US), (Proprietor designated states: all)

INVENTOR:

Addison, Edwin R. Conquest Software Inc., 9700 Patuxent Woods Drive, Suite
140,, Columbia, Maryland MD-21046, (US)

Blair, Arden S. Conquest Software Inc., 9700 Patuxent Woods Drive, Suite
140,, Columbia, Maryland MD-21046, (US)

Nelson, Paul E. Conquest Software Inc., 9700 Patuxent Woods Drive, Suite
140,, Columbia, Maryland MD-21046, (US)

Schwartz, Thomas Conquest Software Inc., 9700 Patuxent Woods Drive, Suite
140, Columbia, Maryland MD-21046, (US)

LEGAL REPRESENTATIVE:

Goodman, Christopher (31122), Eric Potter Clarkson, Park View House, 58
The Ropewalk, Nottingham NG1 5DD, (GB)

PATENT (CC, No, Kind, Date): EP 597630 A1 940518 (Basic)
EP 597630 B1 020731

APPLICATION (CC, No, Date): EP 93308829 931104;

PRIORITY (CC, No, Date): US 970718 921104

DESIGNATED STATES: AT; BE; CH; DE; DK; ES; FR; GB; GR; IE; IT; LI; LU; MC;
NL; PT; SE

INTERNATIONAL PATENT CLASS (V7): G06F-017/27; G06F-017/30

ABSTRACT WORD COUNT: 168

NOTE:

Figure number on first page: 1

LANGUAGE (Publication, Procedural, Application): English; English; English

FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS B	(English)	200231	1139
CLAIMS B	(German)	200231	1201
CLAIMS B	(French)	200231	1291
SPEC B	(English)	200231	11289

Total word count - document A 0

Total word count - document B 14920

Total word count - documents A + B 14920

...ABSTRACT index of "word senses" rather than just words.) It builds its
concept index from a " semantic network " of word relationships with
word definitions drawn from one or more standard human-language
dictionaries...

...SPECIFICATION activation" to identify the meaning of a word in a small
text. Charniak employs a " semantic network " and begins with all
instances of a given word. It then "fans out" in the...

...technique suffers from 2 admitted drawbacks: it requires a high-quality
partially hand-crafted, small semantic network , and this semantic
network is not derived from published sources. Consequently, the
Charniak method has never been applied to...available through the use of
statistical processing on machine readable dictionaries and automatic

acquisition of semantic networks .

Lexical Acquisition

In the field of lexical acquisition, most of the prior art is succinctly...

...for use in natural language processing. None of these proposed the automatic building of a semantic network from published dictionaries.

Indexing

Typical text search systems contain an index of words with references ...into phrases. Their processing is subsumed by the present invention, with the conceptual processing and semantic networks .

Hypertext

Prior art electronically-retrieved documents use "hypertext", a form of manually pre-established cross...

...the abstract is based upon concepts, not just keywords. In addition, the present invention uses semantic networks to further abstract these concepts to gain some general idea of the intent of the...

...or three ranking criteria. No known system in the prior art is capable of acquiring semantic network information directly from published dictionaries, and thus, to the extent that such networks are used...a document, retrieve similar documents;

Private Concept: define a new term, enter it in the " semantic network ", search.

The method of the present invention continues to provide Boolean and statistical query options...

...word or idiom. The method of the present invention builds its concept index from a " semantic network " of word relationships with word definitions drawn from one or more standard English dictionaries. During ...a semantic "word sense disambiguation" and takes place via a "spreading activation" concept through a " semantic network ". The method used disambiguates word senses (identify word meanings) based on "concept collocation". If a...

...to recent words in the text Hence, recent syntactically compatible terms are compared through the semantic network (discussed below) by " semantic distance". A classic example is that the word "bank" when used in close proximity to...

...when used in close proximity to "check".

To make this concept work correctly, an underlying semantic network defined over the word senses is needed. An example of such a network is illustrated...

...0 to 1. Past industrial experience with commercial systems has shown difficulty in maintaining rich semantic networks with many link types. Further, this concept indexing scheme does not require a deep understanding...

...in a local region about the word in question and compared against terms in a " semantic network " that is derived directly from published dictionaries (see discussion below on automatic acquisition.) The resulting...

...breakout of the concept indexing process. The process extracts sentences from the text, tags the words within those sentences, looks up words and analyzes morphology, executes a robust syntactic parse, disambiguates word senses and produces the index.

The first step in the indexing process is to extract sentences or other ...of the word, b) The parts of speech for each meaning, c) Pointers into the semantic networks for each meaning, and d) Information on how the word is used in idioms.

As...

...They are represented as follows:

1. noun, "round spherical object" Word Sense A9C2 (pointer into semantic network)
2. verb, "to gather into a ball, wad" Word Sense A9C3
3. noun, "dance or...

...attempted in parsing. However, when sentences are ungrammatical or unwieldy, or when the input text string is not a full sentence, the chart parser will produce phrase or fragment parses. Hence, the output of the parser may be a...

...a semantic word sense disambiguation and takes place via a spreading activation concept through a semantic network. Figure 3 illustrates the concept which is to disambiguate word senses based on "concept collocation...

...to recent words in the text. Hence, recent syntactically compatible terms are compared through the semantic network by spreading activation or semantic "distance".

An underlying semantic network defined over the word senses is used in this step.. Note that only an "association ...than directly from hardware to software.

The relationship between tool-1 and software, both significant words in the parse of the sentence, has weight 0.35. By observing Figure 5, note that the relationship between tool-2...

...requires no generation of semantic interpretation rules for each word sense. Instead, it requires a semantic network. A later section defines how the method of the present invention acquires the required semantic network by automated means. A key claim in this invention is the use of underlying publisher's dictionaries to produce semantic networks combined with word sense disambiguation, as used here.

The fifth and final step in the...

...group. Each indicator answers a particular question useful for text retrieval, such as "does any document in this group contain the word 'X'? Besides the mere presence or absence of a word in any document, indicators may...

...art systems could only rank documents after the entire search was complete.

Automatic Acquisition of Semantic Networks

One or more publisher's dictionaries (in machine-readable form) may be loaded into a "semantic network", see Figure 4. This is a network of word meanings and relationships to other word...

...database specific terms, idioms or acronyms, by scanning text for concepts not already in the semantic network and by adding them by heuristic association. Finally, non-dictionary data may be added to the

semantic network , such as almanac data or business names and SIC number listings. This enables the retrieval...

...Princeton's Word-Net (George Miller of Princeton University, has produced a 60,000 term semantic network of basic English terms). A benefit of this method is the ability to add or...

...dictionary provides access functions to allow these algorithms to operate. In addition, Princeton's "Word- Net ", a semantic net of English word senses, is used as a machine readable source.

The Composite Dictionary Figure...

...relationships as specified in the thesaurus. Typically, a thesaurus will specify the meaning of the word which contains the relationships .

* Semantic network links - The WordNet format (from Princeton University) is a semantic network of words which links meanings of words to "concepts" (AKA "synonym sets") which are linked...

...word are merged. The "closeness" of two meanings can be determined by looking into the semantic network and computing a distance factor based on the number and the weight of links required...

...operating on large bodies of text may be used to acquire additional dictionary words and semantic network nodes and links. These tools include the following:

1. Find missing words: A dictionary and...horizontal bar) Using Syntactic and Semantic Information

The method of the present invention uses its semantic network to "explode" queries into related concepts and then to "paraphrase" the queries into many different...then takes place by adding closely related word senses extracted via spreading activation from the semantic word sense network .

The augmented query is then used to reference the concept index and the document reference...

...close concept in the index based upon the closeness of the word sense in the semantic word sense network , the syntactic position relative to the query, the modifiers used in association with the head...by constructing relationships in a conceptual graph. This subject is then attached to the underlying semantic network . The concept or topic may then be searched for by using it within a plain...

...file according to a predefined specification. This conceptual graph then gets attached to the underlying semantic network . Each relationship type (not necessarily each individual link) has had a predetermined link strength from...

...include all of their related concepts. This is done by using spreading activation with the semantic network .

- 3) Determine the most frequent concepts in the document, using histograms or some other technique...

...CLAIMS term; the parts of speech of each such meaning; pointer data structures into an associated semantic network for each such meaning; and information about the use of the term in linguistic idioms...

...linguistic databases, suitable for determining one or more likely meanings of identified terms in a query , each term being

identified as a word , phrase or sentence, characterised by the steps of:

(a) identifying root words and their associated meanings...

?

16/3,K/1 (Item 1 from file: 348)
DIALOG(R)File 348:EUROPEAN PATENTS
(c) 2006 European Patent Office. All rts. reserv.

00810047

Method and apparatus for generating query responses in a computer-based document retrieval system

Verfahren und Gerat, um Suchantworten in einem rechnergestutzten Dokumentwiederauffindungssystem zu generieren

Procede et dispositif pour generer des reponses de recherche dans un systeme de recouvrement de documents base sur ordinateur

PATENT ASSIGNEE:

SUN MICROSYSTEMS, INC., (1392730), 2550 Garcia Avenue, Mountain View, CA 94043, (US), (Proprietor designated states: all)

INVENTOR:

Woods, William A., 41 Wildwood Street, Winchester, Massachusetts 01890, (US)

LEGAL REPRESENTATIVE:

Harris, Ian Richard (72231), D. Young & Co., 21 New Fetter Lane, London EC4A 1DA, (GB)

PATENT (CC, No, Kind, Date): EP 752676 A2 970108 (Basic)
EP 752676 A3 980617
EP 752676 B1 021127

APPLICATION (CC, No, Date): EP 96305010 960705;

PRIORITY (CC, No, Date): US 499268 950707

DESIGNATED STATES: DE; FR; GB; NL; SE

INTERNATIONAL PATENT CLASS (V7): G06F-017/30

ABSTRACT WORD COUNT: 183

NOTE:

Figure number on first page: 2

LANGUAGE (Publication,Procedural,Application): English; English; English

FULLTEXT AVAILABILITY:

Available Text	Language	Update	Word Count
CLAIMS B	(English)	200248	652
CLAIMS B	(German)	200248	662
CLAIMS B	(French)	200248	719
SPEC B	(English)	200248	11264
Total word count - document A			0
Total word count - document B			13297
Total word count - documents A + B			13297

...SPECIFICATION out, and a term occurrence index 80 comprising an index of all, or some specified subset of, the terms within the document corpus, as described in further detail below. In addition, generator store 85 is a portion...

...operation is a conventional operation in information retrieval.

The terminology analysis module 100 analyzes each term in the corpus 70 to construct the term /concept relationship network 110, which is a corpus-specific semantic network of terms and concepts that occur in the corpus 70, or related terms and concepts...

...terms and concepts that may be used subsequently to connect terms in a query with terms in the text.

The construction of the term /concept relationship network 110 draws upon and makes use of a lexicon 180 composed of a general...

...words that are used by morphological analysis routines within the terminology analysis module 100 to derive morphological relationships between terms that may not occur explicitly in the lexicon. The

operation and use of such lexicons and morphological analysis is conventional in computational linguistics.

The construction of the term /concept relationship network 110 also makes use of a taxonomy 120 composed of a general purpose taxonomy ...to the subject domain of the corpus 70. This operation also makes use of a semantic network of semantic entailment relationships 150 composed of a general purpose entailments database 160 of semantic entailment relationships...

...specific to the subject domain of the corpus 70. The operation and use of such semantic taxonomies and semantic networks are conventional in the art of knowledge representation. See John Sowa (ed.), Principles of Semantic Networks : Explorations in the Representation of Knowledge, San Mateo: Morgan Kaufmann, 1991.

Each of these modules...combination is computed from the distances between the individual term hits, the similarity distances or match penalties involved in each of the term hits, syntactic information about the region of the hit passage (such as whether there is a sentence ...

...string of the retrieved passage. The retrieved passage is determined by starting with the latest sentence or segment boundary in the source document that precedes the earliest term hit in this match and ends at the first sentence or segment boundary that follows the latest term hit.

6. The displayed term hit list can be used to access a display of... turn can be related to each other by the following morphological, taxonomic, and semantic entailment relationships :

1. term x is a root form of an inflected or derived term y.
2. term or concept x taxonomically subsumes term or concept y (i.e., term...some term occurrence in some file and continuing until some proximity horizon beyond that root term occurrence has been reached.

21. Generating the Term /Concept Relationship Network

During indexing as described in Section 1 above (or in a separate pass) as...

...or phrase in the indexed material is encountered, it is looked up in a growing term /concept relationship network 110 of words and concepts and relationships among them that is being constructed as the corpus is analyzed. If the word or phrase is not already present in...

...time each such word or phrase is encountered, it is also looked up in manually constructed external knowledge bases of word and concept relationships (120, 150 and 180), and if it is found in these external networks, then all words and concepts in the external networks that are known to be entailed by this word or phrase or that are derived or inflected forms of this word or phrase are added to the growing term /concept relationship 110 network together with the known relationships among them. If such a word...

...and 180), and if so, its morphological relationship to its root is recorded in the term /concept relationship network and its root form is treated as if it had occurred in the corpus (i.e., that root is...

...its entailments, inflections, derivations, and relationships are added).

At the end of this process, a term /concept relationship network will have been constructed that contains all of the terms that occur in the corpus plus all of the concepts entailed by or morphologically related...

Set	Items	Description
S1	6588147	SENTENCE? ? OR PARAGRAPH? ? OR DOCUMENT? ? OR (STREAM? ? (3-N)TEXT) OR HYPERTEXT? ? OR WEBPAGE? ? OR HOMEPAGE? ? OR PAGE? ?
S2	12044573	WORD? ? OR TOKEN? ? OR TERM? ? OR STRING? ?
S3	360665	(TOKENI???? OR PARS??? OR ANALY???? OR EVALUAT??? OR EXAMI-N???? OR DIVID??? OR DIVIS??? OR SEPARAT??? OR SPLIT???? OR -GROUP??? OR SECTION??? OR SEGMENT???? OR SUBDIVI???? OR -PARTIT???? OR COMPARTMENT???? OR SUBSET? ? OR SUB() (SET? ? OR TYPE? ? OR
S4	7268	S3 (5N) S2
S5	6124488	PATTERN? ? OR SYNTA???? OR TEMPLATE? ? OR PHRASE? ? OR R-ELATIONSHIP? ?
S6	285602	S2 (3N) S5
S7	61900	(CONSTRUCT??? OR BUILD??? OR CREAT??? OR MAKE OR MAKING OR MADE OR DEVELOP??? OR RECREAT??? OR ASSEMBL??? OR GENERAT??? -OR NEW OR FRESH OR " () SCRATCH OR DERIV???? OR PRODUC??? OR -FORM??? OR SYNTHESI???? OR MANUFACTUR???) (5N) S6 FROM"
S8	61900	S7 (5N) S2
S9	139817	(EACH OR ALL OR EVERY) (2W) S2
S10	4814	(SEARCH??? OR RETRIEV??? OR QUER???? OR MATCH??? OR FIND??? OR TARGET???) (5N) S9
S11	364	S10 (5N) S6
S12	2461	SEMANTIC (3N) (NETWORK? ? OR GRAPH? ? OR TREE? ? OR NET? ?)
S13	0	S11 (5N) S12
S14	0	S11 (10N) S12
S15	4	S11 AND S12
S16	0	S4 AND S8 AND S15
S17	243	S4 AND S8
S18	156	S17 AND (PY<2001 OR PD<20000307)
S19	98	RD (unique items)

? show files

File 275:Gale Group Computer DB(TM) 1983-2006/Sep 11
(c) 2006 The Gale Group

File 47:Gale Group Magazine DB(TM) 1959-2006/Sep 08
(c) 2006 The Gale group

File 16:Gale Group PROMT(R) 1990-2006/Sep 11
(c) 2006 The Gale Group

File 624:McGraw-Hill Publications 1985-2006/Sep 11
(c) 2006 McGraw-Hill Co. Inc

File 484:Periodical Abs Plustext 1986-2006/Sep W1
(c) 2006 ProQuest

File 613:PR Newswire 1999-2006/Sep 11
(c) 2006 PR Newswire Association Inc

File 813:PR Newswire 1987-1999/Apr 30
(c) 1999 PR Newswire Association Inc

File 239:Mathsci 1940-2006/Oct
(c) 2006 American Mathematical Society

File 370:Science 1996-1999/Jul W3
(c) 1999 AAAS

File 696:DIALOG Telecom. Newsletters 1995-2006/Sep 09
(c) 2006 Dialog

File 621:Gale Group New Prod. Annou. (R) 1985-2006/Sep 08
(c) 2006 The Gale Group

File 674:Computer News Fulltext 1989-2006/Aug W4
(c) 2006 IDG Communications

File 88:Gale Group Business A.R.T.S. 1976-2006/Sep 08
(c) 2006 The Gale Group

File 369:New Scientist 1994-2006/Jul W5
(c) 2006 Reed Business Information Ltd.

File 160:Gale Group PROMT(R) 1972-1989

(c) 1999 The Gale Group
File 635:Business Dateline(R) 1985-2006/Sep 09
(c) 2006 ProQuest Info&Learning
File 15:ABI/Inform(R) 1971-2006/Sep 11
(c) 2006 ProQuest Info&Learning
File 9:Business & Industry(R) Jul/1994-2006/Sep 08
(c) 2006 The Gale Group
File 13:BAMP 2006/Sep W1
(c) 2006 The Gale Group
File 810:Business Wire 1986-1999/Feb 28
(c) 1999 Business Wire
File 610:Business Wire 1999-2006/Sep 11
(c) 2006 Business Wire.
File 647:CMP Computer Fulltext 1988-2006/Oct W4
(c) 2006 CMP Media, LLC
File 98:General Sci Abs 1984-2006/Jul
(c) 2006 The HW Wilson Co.
File 148:Gale Group Trade & Industry DB 1976-2006/Sep 11
(c)2006 The Gale Group
File 634:San Jose Mercury Jun 1985-2006/Sep 09
(c) 2006 San Jose Mercury News
File 636:Gale Group Newsletter DB(TM) 1987-2006/Sep 08
(c) 2006 The Gale Group

3914 TOKENI????
 189641 PARS???
 10150426 ANALY????
 3780356 EVALUAT???
 2480665 EXAMIN?????
 2194129 DIVID???
 4595893 DIVIS???
 3492474 SEPARAT???
 1059977 SPLIT????
 14958229 GROUP???
 3503745 SECTION???
 2350297 SEGMENT?????
 222589 SUBDIVI??????
 149240 PARTIT?????
 100128 COMPARTMENT??????
 296598 SUBSET? ?
 977717 SUB
 9358329 SET? ?
 25326845 TYPE? ?
 14854484 GROUP? ?
 16453 SUB(W) ((SET? ? OR TYPE? ?) OR GROUP? ?)
 62612 CLEAV???
 143068 SEGREGAT???
 54477 SUBSECTION???
 6588147 S1
 S3 360665 (TOKENI???? OR PARS??? OR ANALY???? OR EVALUAT??? OR
 EXAMIN????? OR DIVID??? OR DIVIS??? OR SEPARAT??? OR
 SPLIT???? OR GROUP??? OR SECTION??? OR SEGMENT????? OR
 SUBDIVI?????? OR PARTIT????? OR COMPARTMENT?????? OR
 SUBSET? ? OR SUB() (SET? ? OR TYPE? ? OR GROUP? ?) OR
 CLEAV??? OR SEGREGAT??? OR SUBSECTION???) (5N) S1

5185926 CONSTRUCT???
 8916652 BUILD???
 10592348 CREAT???
 12003267 MAKE
 6431863 MAKING
 11060375 MADE
 13007701 DEVELOP???
 385657 RECREAT???
 1791637 ASSEMBL???
 7076585 GENERAT???
 32471975 NEW
 1122516 FRESH
 0 FROM
 222211 SCRATCH
 0 FROM (W) SCRATCH
 1473180 DERIV?????
 27534727 PRODUC???
 13610388 FORM???
 481894 SYNTHESI????
 14345553 MANUFACTUR???
 285602 S6
 S7 61900 (CONSTRUCT??? OR BUILD??? OR CREAT??? OR MAKE OR MAKING
 OR MADE OR DEVELOP??? OR RECREAT??? OR ASSEMBL??? OR
 GENERAT??? OR NEW OR FRESH OR "FROM" () SCRATCH OR
 DERIV????? OR PRODUC??? OR FORM??? OR SYNTHESI???? OR
 MANUFACTUR???) (5N) S6

Set	Items	Description
S1	0	AU=((SIMPSON D? OR SIMPSON, D?) AND (USEY R? OR USEY, R?))
S2	2	AU=(SIMPSON D? OR SIMPSON, D? OR USEY R? OR USEY, R?) AND - SEMANTIC AND (PY<2001 OR PD<20000307)

? show files

File 2:INSPEC 1898-2006/Sep W1
(c) 2006 Institution of Electrical Engineers

File 6:NTIS 1964-2006/Sep W1
(c) 2006 NTIS, Intl Cpyrght All Rights Res

File 8:Ei Compendex(R) 1970-2006/Sep W1
(c) 2006 Elsevier Eng. Info. Inc.

File 34:SciSearch(R) Cited Ref Sci 1990-2006/Sep W1
(c) 2006 The Thomson Corp

File 65:Inside Conferences 1993-2006/Sep 11
(c) 2006 BLDSC all rts. reserv.

File 94:JICST-EPlus 1985-2006/Jun W1
(c)2006 Japan Science and Tech Corp(JST)

File 99:Wilson Appl. Sci & Tech Abs 1983-2006/Jul
(c) 2006 The HW Wilson Co.

File 148:Gale Group Trade & Industry DB 1976-2006/Sep 11
(c)2006 The Gale Group

File 636:Gale Group Newsletter DB(TM) 1987-2006/Sep 08
(c) 2006 The Gale Group